

Sussex Research Online

Educational cultural analysis of language assessment rubrics: a case study of Japanese language at a British University

Article (Published Version)

Winch, Junko (2020) Educational cultural analysis of language assessment rubrics: a case study of Japanese language at a British University. *Journal of Language for Specific Purposes* (7). pp. 7-14. ISSN 2359-8921

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/90288/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

EDUCATIONAL CULTURAL ANALYSIS OF LANGUAGE ASSESSMENT RUBRICS: A CASE STUDY OF JAPANESE LANGUAGE AT A BRITISH UNIVERSITY

Junko Winch

University of Sussex (Sussex Centre for Language Studies), Falmer,
United Kingdom

J.Winch@sussex.ac.uk

Abstract: *The standard of language assessment is considered to be similar within the same country, but it actually varies from institution to institution even within the UK. Rubrics are important for language teachers to assess students' written work, and it also relates to teachers' objective or subjective marking. This paper looks at Japanese assessment criteria in a British STEM university where students study Japanese in the IWLP context. Using two dimensions from Hofstede et al.'s (2010) cultural taxonomy and Hall's (1976) concept of high- and low-context culture, Japanese language rubrics for the written assessment was analysed in 2017. The findings show that the rubrics examined in this study were under the influence of Hofstede et al.'s (2010) collectivist and strong uncertainty avoidance educational culture. The emphasis on the correct use of grammar was observed and also found that language teachers in this institution grade students' written work more objectively using quantitative method. The rubrics includes instructions which enhance the quality of grading consistent and standardise among all language teachers. This process also helps to justify the first markers' awarded marks to the second marker and also the external examiner. Recommendations are given to language teachers and managers who coordinate languages. Language teachers are recommended to inform students whether the focus is accuracy or creativity as this information affects students in working on their assessed work. It is also recommended for managers at language centres to revise periodically the definition of categories to examine if there are any duplication among the rubrics and update them. Incorporating some aspects of rubrics mentioned in this study may enhance the quality of language teachers' grading to be standardised and consistent.*

Keywords: writing assessment; educational culture; higher education; Japanese language teaching; rubrics.

1. Introduction

A rubric is a document which contains the scales and criteria of what is assessed and how scores are given, but it also functions as a standard of assessment constructs who set it. As each language teacher and institution has a different rubrics for what qualifies as good writings compared to poor writings (Erdosy, 2003; Lumley, 2002; Lumley, 2006). Various factors affect to rubrics, such as teacher's rating style, personal characteristics, rating experience and educational background (Lumley & McNamara, 1995; Weigle, 1998). As language teachers' criteria may not necessarily match with that of institution they belong, each institution sets their own rubrics for their language teachers. Rubrics are usually set by the head of the language department or the person who coordinates the languages, which represents the value and the belief of the person who sets the rubrics. Language departments at British universities usually set their own rubrics which are often

based on the Common European Framework of Reference (CEFR) descriptor. They are often shared with students via Virtual Learning Environment. However, it should be noted that the rubrics in this study is not shared with students.

The purpose of this study is to investigate the Japanese language teachers claim that they have more of a workload in marking and grading compared to that of other institutions they are teaching based on the rubrics. It is true that every assessment in every subject is not the same. However, if the workload of marking and grading of assessment of the same language are substantially different, culture may be considered as a possible candidate to explain the differences. According to Jonsson & Svingby (2007), the majority of rubric-related articles discuss the development and benefits of using rubrics but none have looked into cultural investigation on Japanese language rubrics.

This study is guided by following three Research Questions (RQ):

RQ1 Are there any cultural influences in the rubrics on Japanese language writing assessment?

RQ2 Is the assessment of students' written work a subjective process?

RQ3 What are the strengths and weaknesses of this rubrics?

2. Theoretical framework

Three cultural concepts are used as a theoretical framework and explained in three sections for this study. Hofstede et al.'s (2010) Uncertainty Avoidance (UA) and Collectivist dimensions are explained in the first two section. The final section discusses the second concept by Hall's (1976) high-context culture.

Geert Hofstede is considered as one of the leading academics on culture (Kirkman et al., 2006; Merkin et al., 2014). Despite some criticism (e.g., Baskerville, 2003; McSweeney, 2002; Spector, Cooper & Sparks, 2001; Taras & Steel, 2009), 'Hofstede's model has been used most often' (Merkin et al., 2014, p. 3). This study utilises Hofstede's framework as a basis of assessment marking scheme analysis. Hofstede et al.'s (2010) cultural taxonomy consists of five dimensions and has two opposing poles. Having two opposing poles on the spectrum is the main reason to choose as a framework for this study as it is easier to compare and understand educational culture and underlying pedagogies. Categorising a particular nationality into one of two cultures may be too stereotypical and simplistic as the reality is much more complex. Given that today's society consists of people with different heritages and preferences due to globalisation, it is difficult to generalise the cultural preferences of a particular nationality or heritage. However, we cannot dismiss that there is also some truth of Hofstede et al.'s cultural taxonomy.

Hofstede et al. (2010) divide cultures into five dimensions: large vs. small power distance; individualism vs. collectivism; masculinity vs. femininity; strong vs. weak uncertainty avoidance; and long- vs. short-term. Among these, uncertainty avoidance is explained next, which is followed by collectivist and high context culture.

2.1. Strong UA and accuracy

Hofstede et al. (2010) define UA as 'the extent to which the members of a culture feel threatened by ambiguous or unknown situations' (Hofstede et al., 2010, p. 191). High/Strong Uncertainty Avoidance Index (UAI) scoring countries need predictability and low/weak UAI scoring nations are not concerned about unknown situations. Japan is one of high/strong UAI scoring nations (Hofstede et al., 2010) and one of the pedagogies of a strong UA culture include 'precision and punctuality come naturally' (Hofstede, 1991, p. 37). Following Stevens (1998) summarises the different emphasis on reward comparing the British and the Japanese.

The Japanese feel most comfortable in a situation where there is only one correct answer that it is possible to find. They also expect to be rewarded for accuracy. The British, however, expect to be rewarded for originality (Stevens, 1998).

The emphasis on accuracy increases a sense of security for the students, which helps to reduce uncertainty and contributes to strong UA culture. It is hypothesised that the emphasis on accuracy may be found in this study as the rubrics are set by the Japanese coordinator.

2.2. Collectivist culture and information sharing

Collectivist is defined as 'the interests of the individual prevail over the interests of the group' (Hofstede et al., 2010, p. 91). Hofstede (1991) uses Japan as being an example of a collectivist country (Hofstede, 1991, p. 57). It is hypothesised that the writing rubrics set by a Japanese coordinator may have a collectivist cultural influence. Varner & Beamer (2005) explains the relationship between collectivist and information sharing as follows.

Information belongs to the group, not the individual. That way, individuals are linked together into a collective... In group-oriented cultures, what is known by one member of a group is known by all members of the group (Varner & Beamer, 2005, p. 241).

Information sharing is also hypothesised to be found in the writing rubrics.

2.3. Writing style: indirect and vague

Hall's (1976) high-context culture is the second concept. Hall (1976, p. 79) defines 'high-context (HC) communication as 'very little is in the coded, explicit, transmitted part of the message' and 'HC cultures tend to use indirect, non-confrontational, and vague language, relying on the listener's or reader's ability to grasp the meaning from the context' (Hall, 1976, p. 84). Charnock (2010) also points out that: 'Confucian-heritage writers show respect for their readers by presenting material without spelling out its relevance and allowing the reader to draw inference from it' (Charnock, 2010). Both Hall (1976) and Charnock (2010) agree that the Japanese uses vague language and relying on the readers' ability to grasp the meaning. As the rubrics are set by a Japanese coordinator, a high-context (Hall & Hall, 1990) cultural influence may be influenced.

Kaplan (1966) analysed writings of several students from different cultures and he also hypothesised their thought pattern. Kaplan (1966) asserts that 'logic which is the basis of rhetoric, is evolved out of culture; is not universal' and analyses the writing of five different languages (English, Romance, Russian, Oriental and Semitic). Kaplan (1966) describes oriental writing as 'turning and turning in a widening gyre' (Kaplan, 1966, p. 17) and summarises 'the approach of indirection' (Kaplan, 1966, p. 17). Kaplan's (1966) analysis of Japanese writing matches with vagueness claimed by Hall (1976) and Charnock (2010). The rubrics which are set by the Japanese may also have influence of indirect and vague language.

3. Methodology

This study analyses a Japanese language summative assessment writing rubrics for Japanese beginners' level from a British STEM university in 2016/2017.

3.1. Data collection

Data collection was a convenient sampling which the researcher was able to access by teaching at this STEM university. The assessment of this university consists of six pieces of coursework, a final exam and an oral exam during one year. The writing task is part of the final exam which includes grammar, reading and writing skills. The writing task takes open-ended question. The rubrics for the writing assessment are not shared with students. As rubrics were set by a Japanese coordinator, the researcher translated the appropriate Assessed criteria and Scoring criteria into Japanese for the purpose of this study. Assessed criteria has four criteria: 1) Grammar & Structure, 2) Vocabulary, 3) Spelling, and 4) Content & Organisation. The scoring criteria and guide ranges from 0 to 100%. At this university, following six bands are given: over 80%, 70s, 60s, 50s, 40s and below 30%.

3.2. Data analysis

To find out the strong UA, the number of word related to accuracy (i.e. the term 'correct') which appeared in the rubrics was examined and counted as for strong UA. The collectivist culture was examined by looking at instructions in the rubrics which the person who set the rubrics (coordinator) shares information with all language teachers. To identify the high context culture, whether there are any vague or indirectness instructions in the rubrics.

4. Results and discussion

The results of analysis are discussed on three headings strong uncertainty avoidance, collectivist culture and high context culture.

Strong uncertainty avoidance culture: accuracy

The word 'correct' was mentioned total of five times in the assessment criteria of Spelling, Grammar & Structure and Scoring Guide.

In the Spelling Instruction, the word 'accurate' is mentioned once:

'Please check if the students write hiragana and katakan *accurately*' (ICL, 2016). This is the emphasis which shows strong UA.

In the Grammar & Structure instruction, 'correct' was mentioned four times in the following instructions:

1. 'Please check if students have used studied grammar *correctly*' (ICL, 2016) which is an emphasis on correct grammar, which shows preference for strong uncertainty avoidance culture.
2. 'Please write down all grammar points which are used *correctly* at the bottom of the grid paper' (ICL, 2016).
3. Plus points are given for the *correct* use of adverbs such as 'sometimes', 'always', 'very' and 'not very' (ICL, 2016).
4. Plus points are given to the *correct* use of connectors such as 'and' and 'then' (ICL, 2016).

Furthermore, the following Scoring guide instructions emphasise the accuracy regarding the use of grid paper and word count:

The deductions are instructed in the following two occasions:

- 1) when students do not master how to use the grid papers,
- 2) when students do not write the required number of words (ICL, 2016).

These deduction points emphasises accuracy and also imply control in details which shows strong UA.

Collectivist culture: information sharing

Information sharing is common in collectivist culture which is represented by 'what is known by one member of a group is known by all members of the group' (Varner & Beamer, 2005, p. 241). Information sharing is observed by the coordinator's two instructions. The first instruction implies that the more students use learned vocabulary and grammar, the higher their marks on three criteria on 'Content', 'Grammar' and 'Vocabulary'.

'At this level, students have limited vocabulary and grammar so the rich and interesting content may not be apparent. But those who use various learnt vocabulary and grammar points are usually rich and interesting enough... In other words, those who use various learnt vocabulary and grammar points are usually rich and interesting enough. Those who have a few mistakes but do not include various learnt vocabulary and grammar points are not so rich and interesting in content' (ICL, 2016).

In the second instruction, the coordinator shares information which all language teachers are expected to follow.

'Please write down all grammar points which are used *correctly* at the bottom of the grid paper. By doing so, you will see how many grammar points students used and the variety of their usages. This also helps the second markers to mark... It is helpful for the second markers to underline in red where students used incorrect grammar

(but you don't have to correct them as students do not see these scripts)' (ICL, 2016).

This instruction makes the quality of grading/markings consistent and standardised. It also implies the use of quantitative method for justification of the awarded marks. To justify their marks, each language teacher goes through the following three stages:

- 1) The first marker must count the number of grammar and vocabulary mistakes in each student's essay writing and record the number to justify the marks awarded;
- 2) Taking into account the numbers of the highest (Maximum) and the lowest (Minimum) students' number of mistakes, the average number of mistakes is determined by the first marker;
- 3) Based on the maximum, minimum and average number of mistakes, the benchmark is created by the first marker.

This procedure uses descriptive statistics which each teacher has to find the maximum, minimum and average to determine the benchmark. Furthermore, the benchmark alters every time writing assignments are submitted. As this university sets six essay writing assignments in a year, it takes up a large amount of the language teachers' time for grading in addition to their teaching and marking. From the above three stages, it is possible to say that these assessment criteria are not very subjective.

High-Context (HC) culture: vague and indirect

As far as this study's rubrics are concerned, the instructions are very specific ('Please write down all grammar points which are used *correctly* at the bottom of the grid paper' and 'It is helpful for the second markers to underline in red where students used incorrect grammar (but you don't have to correct them as students do not see these scripts)' (ICL, 2016) and in detailed (plus points and deduction points). It is unexpected results to say that the influence of high context culture is very small from the collected data.

5. Conclusions and Recommendations

5.1. Conclusions

A brief review of the three RQs will enable the key conclusions to be summarised in this study. The conclusion considers answering the following questions in this study:

1. Are there any cultural influences in the assessment criteria?

HC, collectivist and strong uncertainty avoidance culture were observed in the assessment criteria. The accuracy related terms were mentioned five times in the instruction, which indicates a strong UA culture. Information sharing was observed in two instructions. The information was the procedure which all language teachers are expected to follow. The information of deduction and plus points were written very specifically and in detail. Indirect instructions were not observed, which implies that the influence of HC culture was not very strong from this data.

2. Is assessment of students' work a subjective process?

Although the rubric does not specifically mention quantitative methods, the actual work which the Japanese language teachers involves with quantitative methods to justify the first markers' awarded marks to the second and external markers. This indicates that assessment of students' work based on these assessment criteria is not very subjective.

3. What is the strength and weakness of these assessment criteria?

The strength of these assessment criteria is the consistency of the grading by standardisation, whereas the weakness is time-consuming to mark and grade, which creates additional work for language teachers.

5.2. Recommendations

Recommendations for this study considers two parties: 1) the managers/directors who coordinate languages; and 2) language teachers.

1) The managers/directors who coordinate languages

It is recommended that managers/directors who coordinate languages should revise the definitions of Categorisation 1 to examine if there are any duplicated categories and update any unclear, ambiguous or inappropriate definitions to enhance the quality of assessment criteria. Although all language teachers refer to the assessment criteria, they may question the validity of their assessment criteria as they are aware that their assessment criteria may not be very clear and useful. It is worth considering incorporating some aspects of the assessment in this paper to standardise and consistent.

2) Language teachers

As for language teachers, it is important to inform students whether the focus of the institution is on accuracy or creativity. This information affects the students' focus to work on their assessment. For example, if the assessment criteria focus on accuracy, students will focus on writing accurately in grammar and vocabulary use. If the assessment criteria focus on creativity, students will write creatively beyond their level and they do not have to concern about making mistakes.

References

- Baskerville, R. F. (2003). Hofstede never studied culture. *Accounting, Organizations and Society*, 28 (1), 1–14.
- Charnock, K. (2010). The right to reticence. *Teaching in Higher Education*. 15(5), pp. 543–552.
- Erdosy, M. U. (2003). Exploring variability in judging writing ability in a second language: A study of four experienced teachers of ESL compositions. TOFL Research Reports, RR-03-17. Princeton, NJ: Educational Testing Service.
- Hall, E. T. (1976). *Beyond Culture*. New York: Anchor Press, Doubleday.
- Hall, E. & Hall, M. 1990. *Understanding cultural differences: Germans, French and Americans*. Yarmouth: Intercultural Press.
- Hofstede, G. (1991). *Cultures and Organizations – Software of the mind*, New York: McGraw-Hill Book Company.
- Hofstede, G., Hofstede, G. J. & Minkov, M. (2010) *Cultures and Organizations – Software of the mind*. New York, NY: Yarmouth: Intercultural Press.

- Imperial College London. (2016). *Japanese Level 1 Marking Scheme (writing) for summative assessment*.
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144.
- Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Language Learning*. Vol. 16, 1–20.
- Kirkman, B. L., K. B. Lowe, & C. B. Gibson. (2006). A quarter century of culture's consequences: A review of empirical research incorporating Hofstede's cultural values framework. *Journal of International Business Studies*, 37(3), 285–320.
<http://dx.doi.org/10.1057/palgrave.jibs.8400202>
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: what do they really mean to the teachers? *Language Testing*, 19(3), 246–276.
- Lumley, T. (2006). *Assessing second language writing: The teacher's perspective*. Frankfurt and Main: Peter Lang.
- Lumley, T. & McNamara, T. F. (1995). Teacher characteristics and teacher bias: Implications for training. *Language Testing*, 12(1), 54–71.
- McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith – A failure of analysis. *Human Relations*, 55(1), 89–118.
- Merkin, R., Taras, V., & Steel, P. (2014). State of the art themes in cross-cultural communication research: A systematic and meta-analytic review. *International Journal of Intercultural Relations*, 38, 1–23. doi.[org/10.1016/j.ijintrel.2013.10.004](https://doi.org/10.1016/j.ijintrel.2013.10.004)
- Spector, P. E., Cooper, C. L., & Sparks, K. (2001). An international study of the psychometric properties of the Hofstede Values Survey Module 1994: A comparison of individual and country/province level results. *Applied Psychology: An International Review*, 50(2), 269–281.
- Stevens, B. (1998). *An investigation of problems experienced by Japanese students, learning English in Great Britain*. MA dissertation: University of Southampton.
- Taras, V., & Steel, P. (2009). Beyond Hofstede: Challenging the ten testaments of cross-cultural research. In C. Nakata (Ed.), *Beyond Hofstede: Culture frameworks for global marketing and management* (pp. 40–61). Chicago, Ill: Macmillan/Palgrave.
- Varner, I. I. & Beamer, L. (2005). *Intercultural Communication in the Global Workplace*. New York: McGraw-Hill.
- Weigle, S. C. (1998). Using FACETS to model teacher training effects. *Language Testing*, 15, 263–287.